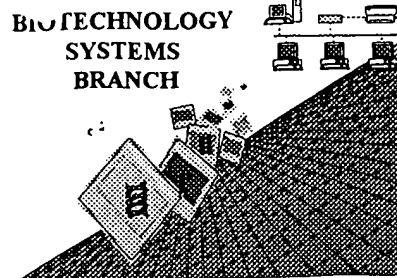


LN

## **RAW SEQUENCE LISTING** **ERROR REPORT**



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/818,954  
Source: OIPE  
Date Processed by STIC: 4/19/2001

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

### **Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO). Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

# Raw Sequence Listing Error Summary

## ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/818,954

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1        Wrapped Nucleics      The number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2        Wrapped Aminos      The amino acid number/text at the end of each line "wrapped" down to the next line.  
This may occur if your file was retrieved in a word processor after creating it.  
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3        Incorrect Line Length      The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4        Misaligned Amino Acid      The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs  
Numbering      between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5        Non-ASCII      This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.  
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6        Variable Length      Sequence(s)        contain n's or Xaa's which represented more than one residue.  
As per the rules, each n or Xaa can only represent a single residue.  
Please present the maximum number of each residue having variable length and  
indicate in the (ix) feature section that some may be missing.
- 7        PatentIn ver. 2.0 "bug"      A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid  
sequence(s)       . Normally, PatentIn would automatically generate this section from the  
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section  
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>  
sections for Artificial or Unknown sequences.
- 8        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence:  
(OLD RULES)      (2) INFORMATION FOR SEQ ID NO:X:  
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")  
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:  
This sequence is intentionally skipped  
  
Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9        Skipped Sequences      Sequence(s)        missing. If intentional, please use the following format for each skipped sequence.  
(NEW RULES)      <210> sequence id number  
<400> sequence id number  
000
- 10        Use of n's or Xaa's      Use of n's and/or Xaa's have been detected in the Sequence Listing.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if n's or Xaa's are present.  
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11        Use of "Artificial"      Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules.  
(NEW RULES)      Valid response is Artificial Sequence.
- 12        Use of <220>Feature      Sequence(s)        are missing the <220>Feature and associated headings.  
(NEW RULES)      Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial Sequence" or "Unknown"  
Please explain source of genetic material in <220> to <223> section.  
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13        PatentIn ver. 2.0 "bug"      Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted  
file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).  
Instead, please use "File Manager" or any other means to copy file to floppy disk.

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/818,954

DATE: 04/19/2001  
TIME: 13:37:55

Input Set : A:\A-676B.ST25.txt  
Output Set: N:\CRF3\04192001\I818954.raw

Does Not Comply  
Corrected Diskette Needed

*pr 4-5*

3 <110> APPLICANT: Paszty, Christopher  
4 Cao, Jin  
5 Danilenko, Dmitry  
6 Gong, Jianhua  
7 Hill, David  
9 <120> TITLE OF INVENTION: Beta-Like Glycoprotein Hormone Polypeptide and Heterdimer  
11 <130> FILE REFERENCE: A-676B  
13 <140> CURRENT APPLICATION NUMBER: US/09/818,954  
14 <141> CURRENT FILING DATE: 2001-03-27  
16 <150> PRIOR APPLICATION NUMBER: 09/723,970  
17 <151> PRIOR FILING DATE: 2000-11-27  
19 <150> PRIOR APPLICATION NUMBER: 60/199,211  
20 <151> PRIOR FILING DATE: 2000-04-24  
22 <150> PRIOR APPLICATION NUMBER: 60/192,654  
23 <151> PRIOR FILING DATE: 2000-03-28  
25 <160> NUMBER OF SEQ ID NOS: 28  
27 <170> SOFTWARE: PatentIn version 3.0  
29 <210> SEQ ID NO: 1  
30 <211> LENGTH: 130  
31 <212> TYPE: PRT  
32 <213> ORGANISM: Homo sapiens  
34 <400> SEQUENCE: 1  
36 Met Lys Leu Ala Phe Leu Phe Leu Gly Pro Met Ala Leu Leu Leu Leu  
37 1 5 10 15  
39 Ala Gly Tyr Gly Cys Val Leu Gly Ala Ser Ser Gly Asn Leu Arg Thr  
40 20 25 30  
42 Phe Val Gly Cys Ala Val Arg Glu Phe Thr Phe Leu Ala Lys Lys Pro  
43 35 40 45  
45 Gly Cys Arg Gly Leu Arg Ile Thr Thr Asp Ala Cys Trp Gly Arg Cys  
46 50 55 60  
48 Glu Thr Trp Glu Lys Pro Ile Leu Glu Pro Pro Tyr Ile Glu Ala His  
49 65 70 75 80  
51 His Arg Val Cys Thr Tyr Asn Glu Thr Lys Gln Val Thr Val Lys Leu  
52 85 90 95  
54 Pro Asn Cys Ala Pro Gly Val Asp Pro Phe Tyr Thr Tyr Pro Val Ala  
55 100 105 110  
57 Ile Arg Cys Asp Cys Gly Ala Cys Ser Thr Ala Thr Thr Glu Cys Glu  
58 115 120 125  
60 Thr Ile  
61 130  
63 <210> SEQ ID NO: 2  
64 <211> LENGTH: 390  
65 <212> TYPE: DNA  
66 <213> ORGANISM: Homo sapiens  
68 <400> SEQUENCE: 2  
69 atgaagctgg cattcctctt ccttggtccc atggccctcc tccttctggc tggetatggc 60  
71 tgtgtcctcg gtgcctccag tgggaacctg cgcaccttg tgggctgtgc cgtgagggag 120

## RAW SEQUENCE LISTING

DATE: 04/19/2001

PATENT APPLICATION: US/09/818,954

TIME: 13:37:55

Input Set : A:\A-676B.ST25.txt

Output Set: N:\CRF3\04192001\I818954.raw

```

73 tttactttcc tggccaagaa gccaggctgc aggggccttc ggatcaccac ggatgcctgc 180
75 tggggctcgt gtgagacctg ggagaaaccc attctggaac cccctatat tgaagcccat 240
77 catcgagtct gtacctacaa cgagacaaaa caggtgactg tcaagctgcc caactgtgcc 300
79 ccgggagtcg accccttcta cacctatccc gtggccatcc gctgtgactg cggagcctgc 360
81 tcactgcca ccacggagtg tgagaccatc 390
84 <210> SEQ ID NO: 3
85 <211> LENGTH: 106
86 <212> TYPE: PRT
87 <213> ORGANISM: Homo sapiens
89 <400> SEQUENCE: 3
91 Ala Ser Ser Gly Asn Leu Arg Thr Phe Val Gly Cys Ala Val Arg Glu
92 1 5 10 15
94 Phe Thr Phe Leu Ala Lys Lys Pro Gly Cys Arg Gly Leu Arg Ile Thr
95 20 25 30
97 Thr Asp Ala Cys Trp Gly Arg Cys Glu Thr Trp Glu Lys Pro Ile Leu
98 35 40 45
100 Glu Pro Pro Tyr Ile Glu Ala His His Arg Val Cys Thr Tyr Asn Glu
101 50 55 60
103 Thr Lys Gln Val Thr Val Lys Leu Pro Asn Cys Ala Pro Gly Val Asp
104 65 70 75 80
106 Pro Phe Tyr Thr Tyr Pro Val Ala Ile Arg Cys Asp Cys Gly Ala Cys
107 85 90 95
109 Ser Thr Ala Thr Thr Glu Cys Glu Thr Ile
110 100 105
112 <210> SEQ ID NO: 4
113 <211> LENGTH: 24
114 <212> TYPE: DNA
115 <213> ORGANISM: Homo sapiens
117 <400> SEQUENCE: 4
118 atgaagctgg cattcctctt cctt 24
121 <210> SEQ ID NO: 5
122 <211> LENGTH: 21
123 <212> TYPE: DNA
124 <213> ORGANISM: Homo sapiens
126 <400> SEQUENCE: 5
127 gcatgtgctg ctacacagg t 21
130 <210> SEQ ID NO: 6
131 <211> LENGTH: 19
132 <212> TYPE: DNA
133 <213> ORGANISM: Homo sapiens
135 <400> SEQUENCE: 6
136 ctgcaggtgc cttcggatc 19
139 <210> SEQ ID NO: 7
140 <211> LENGTH: 22
141 <212> TYPE: DNA
142 <213> ORGANISM: Homo sapiens
144 <400> SEQUENCE: 7
145 gagacatctc cccactgtgt tt 22
148 <210> SEQ ID NO: 8

```

## RAW SEQUENCE LISTING

DATE: 04/19/2001

PATENT APPLICATION: US/09/818,954

TIME: 13:37:55

Input Set : A:\A-676B.ST25.txt

Output Set: N:\CRF3\04192001\I818954.raw

```

149 <211> LENGTH: 22
150 <212> TYPE: DNA
151 <213> ORGANISM: Homo sapiens
153 <400> SEQUENCE: 8
154 gtttccccca acagaatgtc aa                22
157 <210> SEQ ID NO: 9
158 <211> LENGTH: 23
159 <212> TYPE: DNA
160 <213> ORGANISM: Homo sapiens
162 <400> SEQUENCE: 9
163 atgcctatgg cgtcccctca aac                23
166 <210> SEQ ID NO: 10
167 <211> LENGTH: 28
168 <212> TYPE: DNA
169 <213> ORGANISM: Homo sapiens
171 <400> SEQUENCE: 10
172 ctagtagcga gagaggcgac acatgtca          28
175 <210> SEQ ID NO: 11
176 <211> LENGTH: 130
177 <212> TYPE: PRT
178 <213> ORGANISM: Mus musculus
180 <400> SEQUENCE: 11
182 Met Lys Leu Val Tyr Leu Val Leu Gly Ala Val Ala Leu Leu Leu Leu
183 1                    5                    10                    15
185 Gly Gly Pro Asp Ser Val Leu Ser Ser Ser Gly Asn Leu His Thr
186                    20                    25                    30
188 Phe Val Gly Cys Ala Val Arg Glu Phe Thr Phe Met Ala Lys Lys Pro
189                    35                    40                    45
191 Gly Cys Arg Gly Leu Arg Ile Thr Thr Asp Ala Cys Trp Gly Arg Cys
192                    50                    55                    60
194 Glu Thr Trp Glu Lys Pro Ile Leu Glu Pro Pro Tyr Ile Glu Ala Tyr
195 65                    70                    75                    80
197 His Arg Val Cys Thr Tyr Asn Glu Thr Arg Gln Val Thr Val Lys Leu
198                    85                    90                    95
200 Pro Asn Cys Ala Pro Gly Val Asp Pro Phe Tyr Thr Tyr Pro Met Ala
201                    100                   105                   110
203 Val Arg Cys Asp Cys Gly Ala Cys Ser Thr Ala Thr Thr Glu Cys Glu
204                    115                   120                   125
206 Thr Ile
207                    130
209 <210> SEQ ID NO: 12
210 <211> LENGTH: 393
211 <212> TYPE: DNA
212 <213> ORGANISM: Mus musculus
214 <400> SEQUENCE: 12
215 atgaagttgg tataccttgt ccttggtgca gtggccctcc ttctcctggg tggccctgac        60
217 tctgtcctca gcagctccag tgggaacctg cacacttttg tgggctgtgc tgtgagggaa       120
219 ttcactttca tggccaagaa gccaggctgc aggggacttc ggatcaccac agatgcctgc       180
221 tggggccgct gcgagacctg ggagaaaccc atcctggagc ctccctacat tgaagcctat       240

```

## RAW SEQUENCE LISTING

DATE: 04/19/2001

PATENT APPLICATION: US/09/818,954

TIME: 13:37:55

Input Set : A:\A-676B.ST25.txt

Output Set: N:\CRF3\04192001\I818954.raw

```

223 catcgagtgt gtacatacaa tgagaccaga caggtgacag tgaagctgcc taactgtgcc 300
225 cctggagtcg atcctttcta cacctaccct atggctgtcc gatgtgactg tggggcgtgt 360
227 tccactgcca ccaactgagt tgagaccatc tga 393
230 <210> SEQ ID NO: 13
231 <211> LENGTH: 106
232 <212> TYPE: PRT
233 <213> ORGANISM: Mus musculus
235 <400> SEQUENCE: 13
237 Ser Ser Ser Gly Asn Leu His Thr Phe Val Gly Cys Ala Val Arg Glu
238 1 5 10 15
240 Phe Thr Phe Met Ala Lys Lys Pro Gly Cys Arg Gly Leu Arg Ile Thr
241 20 25 30
243 Thr Asp Ala Cys Trp Gly Arg Cys Glu Thr Trp Glu Lys Pro Ile Leu
244 35 40 45
246 Glu Pro Pro Tyr Ile Glu Ala Tyr His Arg Val Cys Thr Tyr Asn Glu
247 50 55 60
249 Thr Arg Gln Val Thr Val Lys Leu Pro Asn Cys Ala Pro Gly Val Asp
250 65 70 75 80
252 Pro Phe Tyr Thr Tyr Pro Met Ala Val Arg Cys Asp Cys Gly Ala Cys
253 85 90 95
255 Ser Thr Ala Thr Thr Glu Cys Glu Thr Ile
256 100 105
258 <210> SEQ ID NO: 14
259 <211> LENGTH: 39
260 <212> TYPE: DNA
C--> 261 <213> ORGANISM: Artificial see item 11 on Error Summary sheet
263 <220> FEATURE:
264 <223> OTHER INFORMATION: Oligonucleotide
266 <400> SEQUENCE: 14
267 attactagtt ccaccatgaa gttggtatac cttgtcctt 39
270 <210> SEQ ID NO: 15
271 <211> LENGTH: 36
272 <212> TYPE: DNA
C--> 273 <213> ORGANISM: Artificial item 11
275 <220> FEATURE:
276 <223> OTHER INFORMATION: Oligonucleotide
278 <400> SEQUENCE: 15
279 ttaataatcg atcgtcagat ggtctcacac tcagt 36
282 <210> SEQ ID NO: 16
283 <211> LENGTH: 36
284 <212> TYPE: DNA
C--> 285 <213> ORGANISM: Artificial
287 <220> FEATURE:
288 <223> OTHER INFORMATION: Oligonucleotide
290 <400> SEQUENCE: 16
291 ccgcactagt tccaccatgc ccatggcacc acgagt 36
294 <210> SEQ ID NO: 17
295 <211> LENGTH: 41
296 <212> TYPE: DNA

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/818,954

DATE: 04/19/2001

TIME: 13:37:55

Input Set : A:\A-676B.ST25.txt

Output Set: N:\CRF3\04192001\I818954.raw

```

C--> 297 <213> ORGANISM: Artificial
299 <220> FEATURE:
300 <223> OTHER INFORMATION: Oligonucleotide
302 <400> SEQUENCE: 17
303 gcggcggttcg atcgctagta gcgggagaaa cggcacatat c 41
306 <210> SEQ ID NO: 18
307 <211> LENGTH: 815
308 <212> TYPE: DNA
309 <213> ORGANISM: Mus musculus
311 <400> SEQUENCE: 18
312 atgcccatgg caccacgagt cttgctcctt tgccctgctgg gcctggcagt cactgaaggg 60
314 catagcccag agacagccat cccaggctgc cacttgcaac gtgagtaact ctgcttgggg 120
316 agcggatgga cgggtaaccg gccagcacg gccttcaccg gctgctccct tctctgcttc 180
318 cagccttcaa tgtgacggtg cgcaagtatc gcctcggcac ttgccagggc tcccacgtgg 240
320 cacaggcctg tgtaggacac tgtgagtcta gtgctttccc ttcccgggtac tctgtgctgg 300
322 tggccagtgg ctatcggcac aacatcacct cttcctccca gtgctgcacc atcagcagcc 360
324 tcagaaaggt aaggggcctg agcctgatgg agcgtgaggg tggggaccca ggggcctgag 420
326 cctgatggag cgtgaggggtg gggacccagg ggtccgaacc tgacctggtg tgaggggtggg 480
328 gacccaggag cccgaacctg accaggtatg agggtgggga cccagggggc cgaacctgac 540
330 cgggggtgtaa gggtggggtc cccaggggac ccgaacctga ccggggccata agggtgggga 600
332 cccccagggg cccgaacctg accaggtgtg agggtgagga cccagggggt cgaacctgat 660
334 gggggcgtag gggtggggtg gaatgggaac aaacttgggt cctcctccaa cagggtgaggg 720
336 tgtggctgca gtgcgtgggg aaccagcgtg gggagcttga gatctttact gcaagggcct 780
338 gccagtgtga tatgtgccgt ttctcccgct actag 815
341 <210> SEQ ID NO: 19
342 <211> LENGTH: 21
343 <212> TYPE: DNA
344 <213> ORGANISM: Homo sapiens
346 <400> SEQUENCE: 19
347 gcctctagaa agagctggga c 21
350 <210> SEQ ID NO: 20
351 <211> LENGTH: 21
352 <212> TYPE: DNA
353 <213> ORGANISM: Homo sapiens
355 <400> SEQUENCE: 20
356 cgccgtgttc catttatgag c 21
359 <210> SEQ ID NO: 21
360 <211> LENGTH: 39
361 <212> TYPE: DNA
C--> 362 <213> ORGANISM: Artificial
364 <220> FEATURE:
365 <223> OTHER INFORMATION: Oligonucleotide
367 <400> SEQUENCE: 21
368 attactagtt ccaccatgaa gttggtatac cttgtcctt 39
371 <210> SEQ ID NO: 22
372 <211> LENGTH: 36
373 <212> TYPE: DNA
C--> 374 <213> ORGANISM: Artificial
376 <220> FEATURE:

```

*This error also appears in sequence 28.*

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/818,954

DATE: 04/19/2001

TIME: 13:37:56

Input Set : A:\A-676B.ST25.txt

Output Set: N:\CRF3\04192001\I818954.raw

L:13 M:270 C: Current Application Number differs, Replaced Current Application Number  
L:261 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:14  
L:273 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:15  
L:285 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:16  
L:297 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:17  
L:362 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:21  
L:374 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:22  
L:529 M:220 C: Keyword misspelled or invalid format, <213> ORGANISM for SEQ ID#:28